Conditional α-diversity for exchangeable Gibbs partitions driven by the stable subordinator*

Annalisa Cerquetti[†]
Department of Methods and Models for Economics, Territory and Finance
Sapienza University of Rome, Italy

January 20, 2013

Abstract

Asymptotic behaviour of conditional α diversity for the two-parameter Poisson-Dirichlet partition model and for the normalized generalized Gamma model has been recently investigated in Favaro et al. (2009, 2011) with a view to possible applications in Bayesian treatment of species richness estimation. Here we generalize those results to the larger class of mixed Poisson-Kingman species sampling models driven by the stable subordinator (Pitman, 2003).

1 Introduction

The concept of *α-diversity* for infinite exchangeable random partitions Π of the positive integers, induced by randomly sampling from an almost surely discrete probability measure P, was first introduced in Pitman (2003, cfr. Sect. 6.1 Prop. 13) as the random variable S, with $0 < S < \infty$, such that, almost surely for $n \to \infty$,

$$\frac{K_n}{n^\alpha} \longrightarrow S$$

for K_n the number of blocks in a random partition of Π_n of $[n]$ induced by Π and $\alpha \in (0,1)$. For infinite random partitions induced by a r.p.m P whose ranked atoms follow a Poisson-Kingman distribution $PK(\rho_\alpha, \gamma)$ driven by the Lévy density of the stable subordinator $\rho_\alpha = \alpha\Gamma(1-\alpha)^{-1}x^{-\alpha-1}dx$, for $0 < \alpha < 1$, with γ on $(0,\infty)$ some mixing density, Pitman shows that $S = T^{-\alpha}$ where $T = S^{-1/\alpha}$ is the random total sum of the ranked atoms of P and has density γ. (See Pitman, 2006, for a comprehensive reference on exchangeable random partitions). Recently interest in *conditional α-diversity* has emerged in posterior species richness estimation in a Bayesian nonparametric approach to species sampling problems (cfr. Lijoi et al., 2007, 2008; Cerquetti, 2009). Given $K_n = k$ the number of species in a partition $(n_1, \dots, n_k)$ induced

* *AMS (2000) subject classification.* Primary: 60G58. Secondary: 60G09.
† Corresponding author, SAPIENZA University of Rome, Via del Castro Laurenziano, 9, 00161 Rome, Italy. E-mail: `annalisa.cerquetti@gmail.com`

by a *basic* sample of observed species, *conditional α-diversity* is defined as the random variable $S_\alpha^{n,k}$ such that, almost surely for $m \to \infty$

$$\frac{K_m}{m^\alpha}\Big|(K_n = k) \to S_\alpha^{n,k}$$

for K_m the unknown number of *new* species induced by an additional sample of observations of dimension m.

In particular Favaro et al. (2009, 2011) derive distributional results respectively for the conditional α diversity under *two-parameter* (α, θ) *Poisson-Dirichlet priors* (Pitman and Yor, 1997), which are well known to correspond to r.p.m.s whose ranked atoms follow a $PK(\rho_\alpha, \gamma_{\alpha,\theta})$ distribution for $\gamma_{\alpha,\theta}(t) = \frac{\Gamma(\theta+1)}{\Gamma(\theta/\alpha+1)} t^{-\theta} f_\alpha(t)$ for $0 < \alpha < 1$ and $\theta > -\alpha$, and under *normalized generalized Gamma priors*, whose ranked atoms follow a Poisson-Kingman law with mixing density the exponentially tilted version of the α-stable density (cfr. Cerquetti, 2007; Lijoi et al. 2007a). In Cerquetti (2011) a first alternative derivation for the conditional α diversity of the two-parameter Poisson-Dirichlet family has been obtained by a decomposition approach properly exploiting a characterization of those models in terms of the *deletion of classes* property (cfr. Pitman, 2003; Gnedin et al. (2009)). Here we obtain the general result for the entire class of mixed Poisson-Kingman models driven by the stable subordinator, for an arbitrary mixing density γ on $(0, \infty)$ written as, without loss of generality, $\gamma(t) = h(t) f_\alpha(t)$ for some non negative function h such that γ is a proper density and f_α the density of the stable subordinator. This class includes infinitely many priors, and as from Gnedin and Pitman (2006), induces exchangeable random partitions in Gibbs form of type α i.e. uniquely characterized by exchangeable partition probability function (EPPF) in the following Gibbs product form

$$p(n_1, \dots, n_k) = V_{n,k} \prod_{i=1}^{k} (1-\alpha)_{n_j}, \tag{1}$$

where $(a)_b = a(a+1)\cdots(a+b-1)$ are rising factorials and the $V_{n,k}$ are weights satisying the backward recursive relation $V_{n,k} = (n - k\alpha)V_{n+1,k} + V_{n+1,k+1}$.

2 Main results

The basic result in Gnedin and Pitman (2006, cfr Th. 12) states that the EPPF of each exchangeable Gibbs partition of type α, for $\alpha \in (0, 1)$, arises by mixing the EPPF corresponding to the Poisson-Kingman $PK(\rho_\alpha | T = t)$ model with the density γ identifying the specific $PK(\rho_\alpha, \gamma)$ family, hence

$$p_{\alpha,\gamma}(n_1, \dots, n_k) = \int_0^\infty p_\alpha(n_1, \dots, n_k | t)\gamma(t)dt.$$

The explicit formula for $p_\alpha(n_1, \dots, n_k|t)$ was first obtained in Pitman (2003, cfr. eq. (66)) and is central to our main result.

Theorem 1. Let Π be a $PK(\rho_\alpha, \gamma)$ partition of $\mathbb{N}$ for some $0 < \alpha < 1$ and some mixing

probability distribution γ on $(0,\infty)$. Without loss of generality assume $\gamma(t) = h(t)f_\alpha(t)$. Fix $n \geq 1$ and a partition $(n_1,\ldots,n_k)$ of $[n]$ with k positive box-sizes, then Π has conditional α-diversity $S_{\alpha,h}^{n,k}$ with density

$$f_{n,k}^{h,\alpha}(s) = \frac{h(s^{-1/\alpha})\tilde{g}_{n,k}^{\alpha}(s)}{\mathbb{E}_{n,k}^{\alpha}[h(S^{-1/\alpha})]} \tag{2}$$

for

$$\tilde{g}_{n,k}^{\alpha}(s) = \frac{\Gamma(n)}{\Gamma(n-k\alpha)\Gamma(k)} s^{k-1/\alpha-1}\int_0^1 p^{n-1-k\alpha} f_\alpha((1-p)s^{-1/\alpha})dp \tag{3}$$

the density of the product of independent r.v.s $Y_{\alpha,k} \times [W]^\alpha$, where $Y_{\alpha,k}$ has density

$$g_{\alpha,k\alpha}(y) = \frac{\Gamma(k\alpha+1)}{\Gamma(k+1)} y^k g_\alpha(y) \tag{4}$$

for $g_\alpha(y) = \alpha^{-1}y^{-1-1/\alpha}f_\alpha(y^{-1/\alpha})$, and $W \sim \beta(k\alpha, n-k\alpha)$.

Proof: By the unconditional result in Pitman (2003) recalled in the Introduction, the α-diversity for a general $\gamma(t) = h(t)f_\alpha(t)$ mixing density, is the r.v. $S_{\alpha,h}$ with density

$$\gamma(s^{-1/\alpha}) = h(s^{-1/\alpha})f_\alpha(s^{-1/\alpha})\alpha^{-1}s^{-1/\alpha-1}. \tag{5}$$

As from Pitman (2003, cfr. Prop. 9), given $S_\alpha = s$, the exchangeable partition probability function for a $PK(\rho_\alpha|s)$ model corresponds to

$$p_\alpha(n_1,\ldots,n_k|s^{-1/\alpha}) = \frac{\alpha^k}{\Gamma(n-k\alpha)} s^k[f_\alpha(s^{-1/\alpha})]^{-1}\int_0^1 p^{n-1-k\alpha}f_\alpha((1-p)s^{-1/\alpha})dp\prod_{j=1}^{k}(1-\alpha)_{n_j-1},$$

hence by Bayes' rule

$$f_{S_\alpha,\gamma}(s|n_1,\ldots,n_k) = \frac{p_{\alpha,\gamma}(n_1,\ldots,n_k|s^{-1/\alpha})\gamma(s^{-1/\alpha})}{\int_0^\infty p_{\alpha,\gamma}(n_1,\ldots,n_k|s^{-1/\alpha})\gamma(s^{-1/\alpha})ds}$$

which simplifies to

$$f_{S_{\alpha,h}}(s|K_n = k) = \frac{h(s^{-1/\alpha})s^{k-1/\alpha-1}\int_0^1 p^{n-1-k\alpha}f_\alpha((1-p)s^{-1/\alpha})dp}{\int_0^\infty h(s^{-1/\alpha})s^{k-1/\alpha-1}[\int_0^1 p^{n-1-k\alpha}f_\alpha((1-p)s^{-1/\alpha})dp]ds}$$

and the result is proved. Notice that by definition of mixed $PK(\rho_\alpha, h\times f_\alpha)$ model, the general weights $V_{n,k,h}$ in the EPPF (1) arise as follows

$$V_{n,k,h} = \frac{\alpha^{k-1}}{\Gamma(n-k\alpha)}\int_0^\infty h(s^{-1/\alpha})s^{k-1/\alpha-1}\int_0^1 p^{n-1-k\alpha}f_\alpha((1-p)s^{-1/\alpha})dpds$$

hence the normalizing constant in formula (2) may be obtained through the following relationship (see also Ho et al. 2008, eq. (12))

$$\mathbb{E}_{n,k}^{\alpha}[h(S^{-1/\alpha})] = V_{n,k,h}\frac{\alpha^{1-k}\Gamma(n)}{\Gamma(k)}. \tag{6}$$

Remark 2. The result in Theorem 1. agrees with an analogous result for the conditional distribution $T|K_n = k$ (where T is the random total sum of the ranked atoms of the r..p.m. P) first derived in an unpublished manuscript by Ho et al. (2008, cfr. Eq. (13)), that we received by one of those authors as a personal communication. Their result relies on the r.v.

$$R_{\alpha,(n,k)} = \frac{S_{\alpha,k\alpha}}{\beta(k\alpha, n-k\alpha)}$$

for $S_{\alpha,k\alpha}$ the polynomially tilted stable random variable with density

$$f_{S_{\alpha,k\alpha}}(t) = \frac{\Gamma(k\alpha+1)}{\Gamma(k+1)} t^{-k\alpha} f_\alpha(t).$$

It is an easy task to show that $[R_{\alpha,(n,k)}]^{-\alpha} = [Y_{\alpha,k} \times \beta(k\alpha, n-k\alpha)^\alpha]$.

In what follows we show how some distributional results for the conditional α diversity of specific Poisson-Kingman models already obtained in the literature may be easily derived by the general formula in Theorem 1.

Example 3. [*Two-parameter Poisson-Dirichlet* (α, θ) *partition models*] A first result for the conditional α diversity for the two-parameter Poisson-Dirichlet model (Pitman & Yor, 1997) has been derived in Favaro et al. (2009) in view of Bayesian nonparametric posterior interval estimation for the number of new species in an additional sample in species sampling problems. Those authors rely on mimicking the original proof for the unconditional α diversity in Pitman (2006, Th. 3.8). To apply the general result it is enough to notice that the two-parameter Poisson-Dirichlet (α, θ) for $\theta > -\alpha$ partition model is well-known to correspond (cfr. Pitman, 2003) to a mixed Poisson-Kingman model driven by the stable subordinator with mixing density

$$\gamma_{\alpha,\theta}(t) = h(t) \times f_\alpha(t) = \frac{\Gamma(\theta+1)}{\Gamma(\theta/\alpha+1)} t^{-\theta} f_\alpha(t)$$

with the weights in the Gibbs representation of the EPPF corresponding to

$$V_{n,k}^{\alpha,\theta} = \alpha^{k-1} \frac{\Gamma(\theta/\alpha+k)\Gamma(\theta+1)}{\Gamma(\theta+n)\Gamma(\theta/\alpha+1)}.$$

Hence, by (6), the denominator in (2) corresponds to

$$\mathbb{E}_{n,k}^{\alpha}(h(Z^{-1/\alpha})) = \mathbb{E}_{n,k}^{\alpha}(Z^{\theta/\alpha}) = \frac{\Gamma(\theta/\alpha+k)\Gamma(\theta+1)}{\Gamma(\theta+n)\Gamma(\theta/\alpha+1)} \frac{\Gamma(n)}{\Gamma(k)}$$

and by Theorem 1. the conditional α diversity $Z_{n,k}^{\alpha,\theta}$ has density

$$f_{n,k}^{\alpha,\theta}(z) = \frac{\Gamma(\theta+n)}{\Gamma(n-k\alpha)\Gamma(\theta/\alpha+k)} z^{\theta/\alpha+k-1-1/\alpha} \int_0^1 f_\alpha[(zw^{-\alpha})^{-1/\alpha}](1-w)^{n-k\alpha-1} dw \qquad (7)$$

which corresponds to the θ/α polynomial tilting of $\tilde{g}_{n,k}^{\alpha}$. It is an easy task to verify that this is the density of the product of independent r.v.s $Z_{n,k}^{\alpha,\theta} = Y_{\alpha,\theta/\alpha+k} \times [\beta(\theta+k\alpha, n-k\alpha]^\alpha$, as

already established in Cerquetti (2011) moving from a decomposition approach exploiting the *deletion of classes property* of this model (see Gnedin et al., 2009). In the next Proposition we prove the result actually agrees with the original result in Favaro et al. (2009) expressed in terms of the alternative scale mixture representation $Y_{\alpha,(\theta+n)/\alpha} \times \beta(\theta/\alpha + k, n/\alpha - k)$. Notice that the equivalence in distribution between $R_{\alpha,(n,k)}$ and $S_{\alpha,n}[\beta_{k,n/\alpha-k}]^{-1/\alpha}$ was already stated without proof in Ho et al. (2008, cfr. Prop. 2.1, Eq. (11)).

Proposition 4. Let $H = Y_1 \times X$ for Y_1 and X independent r.v.s, $Y_1 \sim g_{\alpha,(\theta+n)}$ and $X \sim \beta(\theta/\alpha+k, n/\alpha-k)$, then the r.v. $Z_{n,k}^{\alpha,\theta}$ with density (7) and H have the same characteristic function

$$G_{\alpha,\theta}^{n,k}(t) = \sum_{r \geq 0} \frac{(it)^r}{r!} \left(\frac{\theta + k\alpha}{\alpha}\right)_r \frac{1}{(\theta+n)_{r\alpha}}.$$

Proof: First notice that by Proposition 2. in Cerquetti (2011), for $m \to \infty$,

$$\mathbb{E}_{\alpha,\theta}\left(\frac{K_m^r}{m^{r\alpha}}\Big|K_n = k\right) \longrightarrow \left(\frac{\theta+k\alpha}{\alpha}\right)_r \frac{\Gamma(\theta+n)}{\Gamma(\theta+n+r\alpha)}.$$

By the change of variable $zw^{-\alpha} = s$ (7) may be written as

$$f_{n,k}^{\alpha,\theta}(z) = \frac{\Gamma(\theta+n)\Gamma(\theta+k\alpha+1)}{\Gamma(\theta+k\alpha)\Gamma(n-k\alpha)\Gamma((\theta+k\alpha)/\alpha+1)} \frac{1}{\alpha} z^{\theta/\alpha+k-1} \times$$

$$\times \int_z^\infty \alpha^{-1} s^{-1/\alpha-1} f_\alpha(s^{-1/\alpha}) \left(1 - (z/s)^{1/\alpha}\right)^{n-k\alpha-1} ds.$$

Its characteristic function is given by

$$G_{n,k}^{\alpha,\theta}(t) = \frac{\Gamma(\theta+n)\Gamma(\theta+k\alpha+1)}{\Gamma(\theta+k\alpha)\Gamma(n-k\alpha)\Gamma((\theta+k\alpha)/\alpha+1)} \frac{1}{\alpha} \times$$

$$\times \int_0^\infty \exp\{itz\} z^{\theta/\alpha+k-1} \int_z^\infty g_\alpha(s) \left(1 - (z/s)^{1/\alpha}\right)^{n-k\alpha-1} dsdz$$

and may be rewritten as

$$G_{n,k}^{\alpha,\theta}(t) = \frac{\Gamma(\theta+k\alpha+1)}{\Gamma((\theta+k\alpha)/\alpha+1)} \frac{1}{\alpha} \int_z^\infty g_\alpha(s) \times$$

$$\times \int_0^\infty \exp\{itz\} \frac{\Gamma(\theta+n)}{\Gamma(\theta+k\alpha)\Gamma(n-k\alpha)} z^{\theta/\alpha+k-1} \left(1 - (z/s)^{1/\alpha}\right)^{n-k\alpha-1} dzds.$$

By the change of variable $(z/s)^{1/\alpha} = y$, $z = y^\alpha s$, $dz = s\alpha y^{\alpha-1}dy$ yields

$$= \frac{\Gamma(\theta+k\alpha+1)}{\Gamma((\theta+k\alpha)/\alpha+1)} \frac{1}{\alpha} \int_0^\infty g_\alpha(s) \times$$

$$\times \int_0^s e^{ity^\alpha s} \frac{\Gamma(\theta+n)}{\Gamma(\theta+k\alpha)\Gamma(n-k\alpha)} (y^\alpha s)^{\theta/\alpha+k-1} (1-y)^{n-k\alpha-1} s\alpha y^{\alpha-1} dyds$$

and then reduces to

$$= \frac{\Gamma(\theta + k\alpha + 1)}{\Gamma((\theta + k\alpha)/\alpha + 1)} \int_0^\infty s^{\theta/\alpha + k} g_\alpha(s) \int_0^1 e^{ity^\alpha s} \frac{\Gamma(\theta + n)}{\Gamma(\theta + k\alpha)\Gamma(n - k\alpha)} (y)^{\theta + k\alpha - 1} (1 - y)^{n - k\alpha - 1} dyds.$$

Exploiting the known characteristic function of Y^α for $Y \sim Beta(\theta + k\alpha, n - k\alpha)$ we can write

$$= \frac{\Gamma(\theta + k\alpha + 1)}{\Gamma((\theta + k\alpha)/\alpha + 1)} \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \frac{(\theta + k\alpha)_{r\alpha}}{(\theta + n)_{r\alpha}} \int_0^\infty s^{\theta/\alpha + k + r} g_\alpha(s) ds$$

and by (4)

$$= \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \frac{(\theta + k\alpha)_{r\alpha}}{(\theta + n)_{r\alpha}} \frac{\Gamma(\theta + k\alpha + 1)}{\Gamma((\theta + k\alpha)/\alpha + 1)} \frac{\Gamma((\theta + k\alpha + r\alpha)/\alpha + 1)}{\Gamma(\theta + k\alpha + r\alpha + 1)}. \tag{8}$$

By the usual properties of Gamma function the last expression corresponds to

$$= \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \frac{\Gamma(\theta + k\alpha + r\alpha)\Gamma(\theta + n)}{\Gamma(\theta + k\alpha)\Gamma(\theta + n + r\alpha)} \frac{(\theta + k\alpha)\Gamma(\theta + k\alpha)}{\frac{\theta + k\alpha}{\alpha}\Gamma(\frac{\theta + k\alpha}{\alpha})} \frac{\Gamma(\frac{\theta + k\alpha}{\alpha} + r)\frac{\theta + k\alpha + r\alpha}{\alpha}}{(\theta + k\alpha + r\alpha)\Gamma(\theta + k\alpha + r\alpha)}$$

which simplifies to

$$= \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \left(\frac{\theta + k\alpha}{\alpha}\right)_r \frac{1}{(\theta + n)_{r\alpha}} \tag{9}$$

and the conclusion follows by the result in Proposition 2 in Favaro et al. (2009) proving that (9) is the characteristic function of $H = Y_1 \times X$.

Example 5. [*Generalized Gamma partition models*] As from Pitman (2003), generalized Gamma partitions models belong to the Poisson-Kingman family driven by the stable subordinator $PK(\rho_\alpha, \gamma)$ for a mixing density

$$\gamma_{\alpha,\lambda}(t) = \exp\{\psi_\alpha(\lambda) - \lambda t\} f_\alpha(t),$$

where $\psi_\alpha(\lambda) = (2\lambda)^\alpha$ is the Laplace exponent of $f_\alpha(\cdot)$. By an application of (5), after the reparametrization $\lambda = \beta^{1/\alpha}/2$, the *unconditional* α diversity for this model (see also Cerquetti (2007) and Lijoi et al. (2007b)) is given by

$$\gamma_{\alpha,\beta}(s^{-1/\alpha}) = \exp\left\{\beta - \frac{1}{2}\left(\frac{\beta}{\alpha}\right)^{1/\alpha}\right\} f_\alpha(s^{-1/\alpha})\alpha^{-1}s^{-1/\alpha - 1}.$$

To obtain the density of the conditional *posterior* α diversity it is enough to apply formula (2) which provides

$$f_{n,k}^{\exp,\alpha}(s) = \frac{\exp\left\{\beta - \frac{1}{2}\left(\frac{\beta}{s}\right)^{1/\alpha}\right\} \tilde{g}_{n,k}^\alpha(s)}{\mathbb{E}_{n,k}^\alpha\left[\exp\left\{\beta - \frac{1}{2}\left(\frac{\beta}{S}\right)^{1/\alpha}\right\}\right]}.$$

The denominator arises by (6) and the known expression for the $V_{n,k}$ of $PK(\rho_\alpha, \gamma_{\alpha,\beta})$ models as obtained in Eq. (6) in Cerquetti (2007),

$$V_{n,k}^{\alpha,\beta} = \frac{e^\beta 2^n \alpha^k}{\Gamma(n)} \int_0^\infty \lambda^{n-1} \frac{e^{-(\beta^{1/\alpha}+2\lambda)^\alpha}}{(\beta^{1/\alpha}+2\lambda)^{n-k\alpha}} d\lambda, \tag{10}$$

which rewritten in terms of incomplete Gamma functions by the change of variable $(\beta^{1/\alpha} + 2\lambda)^\alpha = x$, $d\lambda = (2\alpha)^{-1} x^{1/\alpha-1} dx$ yields

$$V_{n,k}^{\alpha,\beta} = \frac{e^\beta \alpha^{k-1}}{\Gamma(n)} \sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^i (\beta)^{i/\alpha} \Gamma(k - \frac{i}{\alpha}; \beta).$$

By equation (6)

$$\mathbb{E}_{n,k}^\alpha \left[\exp\left\{ \beta - \frac{1}{2} \left(\frac{\beta}{S} \right)^{1/\alpha} \right\} \right] = \frac{e^\beta}{\Gamma(k)} \sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^i (\beta)^{i/\alpha} \Gamma(k - \frac{i}{\alpha}; \beta)$$

hence by Theorem 1. the conditional α diversity $S_{n,k}^{\alpha,\beta}$ for the generalized Gamma model has density

$$f_{n,k}^{\alpha,\beta}(s) = \frac{\Gamma(k) \exp(2^{-1}(\beta/s)^{1/\alpha}) \tilde{g}_{n,k}^\alpha(s)}{\sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^i (\beta)^{i/\alpha} \Gamma(k - \frac{i}{\alpha}; \beta)}.$$

The result agrees with Favaro et al. (2011, Th. 1.) due to the equivalence in distribution between $Y_{\alpha,n/\alpha} \times \beta(k, n/\alpha - k)$ and $Y_{\alpha,k} \times \beta(k\alpha, n - k\alpha)$ which arises specializing for $\theta = 0$ the result in Proposition 4.

References

Cerquetti, A. (2007) A note on Bayesian nonparametric priors derived from exponentially tilted Poisson-Kingman models. *Stat & Prob Letters*, 77, 18, 1705–1711.

Cerquetti, A. (2009) A Generalized sequential construction of exchangeable Gibbs partitions with application. *Proceedings of S.Co. 2009, September 14-16, Milano, Italy.*

Cerquetti, A. (2011) A decomposition approach to Bayesian nonparametric estimation under two-parameter Poisson-Dirichlet priors. *Proceedings of ASMDA 2011 - Applied Stochastic Models and Data Analysis. Rome, June, 7-10 2011*

Favaro, S., Lijoi, A., Mena, R.H. and Prünster, I. (2009) Bayesian non-parametric inference for species variety with a two-parameter Poisson-Dirichlet process prior. *JRSS-B*, 71, 993-1008.

Favaro, S., Lijoi, A. and Prünster, I. (2011) Asymptotics for a Bayesian nonparametric estimator of species variety. *Bernoulli* (to appear).

GNEDIN, A., HAULK, S. AND PITMAN, J. (2009) Characterizations of exchangeable partitions and random discrete distributions by deletion properties. http://arxiv.org/abs/0909.3642

GNEDIN, A. AND PITMAN, J. (2006) Exchangeable Gibbs partitions and Stirling triangles. *Journal of Mathematical Sciences*, 138, 3, 5674–5685.

HO, M-W, JAMES, L.F. AND LAU, J.W. (2008) Explicit Gibbs Chinese Restaurant Process priors. *Unpublished manuscript*

LIJOI, A., MENA, R.H. AND PRÜNSTER, I. (2007a) Controlling the reinforcement in Bayesian non-parametric mixture models. *JRSS-B*, 69, 4, 715–740.

LIJOI, A., MENA, R.H. AND PRÜNSTER, I. (2007b) Bayesian nonparametric estimation of the probability of discovering new species. *Biometrika*, 94, 769–786.

LIJOI, A., PRÜNSTER, I. AND WALKER, S.G. (2008) Bayesian nonparametric estimator derived from conditional Gibbs structures. *Annals of Applied Probability*, 18, 1519–1547.

PITMAN, J. (2003) Poisson-Kingman partitions. In D.R. Goldstein, editor, *Science and Statistics: A Festschrift for Terry Speed*, volume 40 of Lecture Notes-Monograph Series, pages 1–34. Institute of Mathematical Statistics, Hayward, California.

PITMAN, J. (2006) *Combinatorial Stochastic Processes*. Ecole d'Eté de Probabilité de Saint-Flour XXXII - 2002. Lecture Notes in Mathematics N. 1875, Springer.

PITMAN, J. AND YOR, M. (1997) The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, 25, 855–900.